Ensemble Forecasting of Solar Irradiance in Coastal Southern California

Bengu Ozge Akyurek, Xiaohui Zhong

Jan Kleissl – jkleissl@ucsd.edu

Abstract

A solar forecast system with a focus on marine layer cloud forecasting in coastal Southern California during the summer time was developed. Several configuration of the Weather Research and Forecasting (WRF) model are applied to generate a distribution of forecast results for Global Horizontal Irradiance (GHI). The forecasts are validated against GHI observations form 8 weather stations located along a coastal to inland gradient.

The direct outputs from all numerical weather prediction models including WRF are significantly biased. Postprocessing is required to improve upon simple 24 hours persistence forecasts and for the WRF models this improvement is about 30% depending on the ground station. The forecast performance is not sensitive to which input data (WRF output or past observations of GHI or both) are used. Also including additional WRF outputs aside from GHI shows little benefit. The product developed in this research could serve as an operational day-ahead marine layer solar forecast system.

1 Introduction

Weather is a continuous, data-intensive, multidimensional, dynamic and chaotic process, and these properties make weather forecasting a formidable challenge. There is a wide range of techniques involved in weather forecasting from basic approaches to highly complex computerized models (Maqsood, 2002). Accurate forecasting of solar irradiance is essential for the efficient operation of solar thermal power plants, energy markets, and the widespread implementation of solar photovoltaic technology.

Numerical weather prediction (NWP) is generally the most accurate tool for forecasting solar irradiation several hours in advance (Mathiesen et al., 2011). These methods model the weather numerically and use time integration to forecast the future state of the weather and solar irradiance. Despite computational complexity and intensity, the multitude of parameterizations employed and insufficient grid resolution cause the direct output from these models to be inaccurate.

Machine learning techniques, on the other hand, assume that the complex physical relationships can be mapped to simpler functional relationships between the key variables at much smaller computational cost. One family of methods is the analog method family. It hypothesizes that what will happen tomorrow has already happened in the past. Considering past observations and forecasts, called the ensemble set, analog forecasts predict which date or dates are most similar ("analogous") to the forecast period.

Analog methods can be either used for post-processing of a numerical weather model output or for forecasting directly from historical measurements. Post-processing is the process of taking forecast products of another tool and improving them.

This report describes a new analog based forecasting algorithm called Taylor Expanded Solar Analog Forecasting (TESLA) applied to observations and NWP output from coastal California. In southern California, low-altitude marine layer stratocumulus cloud (MLS) cover is common during April through September mornings. Generally these clouds are optically thick and can reduce solar photovoltaic production by up to 70%. The primary objective is to forecast the burn-off of marine layer clouds.

2 Post-Processing Methodology

Our analog method uses past observations and NWP forecasts as input parameters to calculate Global Horizontal Irradiance (GHI) forecasts. The past observation and NWP forecast set is called the ensemble set. Simply put, TESLA provides a numeric function that transforms the ensemble set into a prediction. The input parameters can be the observations from 24 hours ago or the predictions of an NWP, or virtually anything that may or may not seem relevant to the prediction. TESLA uses the ensemble set to "learn/train" how the past inputs are related to the past actual observations and produce the function that maps this connection between the inputs and the prediction. This learning process automatically filters out irrelevant inputs and adjusts the weights of the relevant inputs to produce the function that minimizes the error in the past ensemble set.

TESLA can also be operated only based on past observations without any NWP input. As an example, if we assume that the only input parameter is the observation from 24 hours before the forecast valid time, x, then TESLA could produce a function (1.1x – 100), which indicates that yesterday's observation should be scaled by 1.1 and 100 subtracted to obtain the forecast for tomorrow. The function has to be produced / trained only once based on the ensemble set and then can be used for any number of predictions. The forecast calculation is then nearly as easy as in the example.

2.1 Theoretical Background

TESLA assumes that there is a relationship between some input parameters and future solar irradiance that we want to predict. The assumption is that this connection exists as an unknown function.

This function is then written as a Taylor series expansion. The order of truncation of the series also determines the order of the forecast (see section 2.2.3). The past ensemble set provides the inputs of the function and the exact output of the function is known from past observations. The only unknowns are the coefficients of the Taylor expansion.

The values of the input parameters in the past are combined into a single matrix. Denoting this matrix by **E**, the resulting past observation vector as **O** and the unknown coefficients as **c**, the relation becomes a simple matrix multiplication of **E c**= **O** (Eq. 1). The unknown coefficients can then be easily found by using Least Squares Matrix Division that minimizes the square of the error efficiently.

If the order is greater than 1, the higher order values of the input parameters are added to **E**. As an example for order 2, **E** will contain the ensemble data set values, the square of the values and the cross multiplication of all values. This is the reason, why the complexity increases exponentially with higher orders. For illustration, an example case with two ensemble parameters “x” and “y” in a second order expansion is shown in Equation 1

Equation 1: Example system of equations in TESLA for order 2. The input variables x, and y are matched to the observations O through constants c_0 through c_5.

$$\begin{vmatrix} 1 & x_1 & y_1 & x_1^2 & y_1^2 & x_1 y_1 \\ 1 & x_2 & y_2 & x_2^2 & y_2^2 & x_2 y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_N & y_N & x_N^2 & y_N^2 & x_N y_N \end{vmatrix} \times \begin{vmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{vmatrix} = \begin{vmatrix} o_1 \\ o_2 \\ \dots \\ o_N \end{vmatrix}$$

In addition, the rank of the matrix **E**, both depends on the size of the ensemble set as well as the number of input parameters. It is easy to accidentally create an under-defined system by selecting a high order or high number of input parameters, which will degrade the performance significantly.

2.2 TESLA Configurations

TESLA has multiple configuration parameters that affect the forecast accuracy, the forecast horizon, and the complexity of the training calculation. This section explains the various configuration parameters and their effects on the forecast product. A summary of the configurations is shown in Table 1. Refer to the following sections for more information.

Table 1: TESLA Configurations.

Input Variables (Section 2.1.2)	All	GHI	GHI	GHI	GHI	GHI	Surface Temperature
		Columnar Cloud Cover			Columnar Cloud Cover		Temperature at 850&700hPa
		Surface Pressure		Surface Pressure			
		The height at which atmospheric pressure 850&700hPa	The height at which atmospheric pressure 700hPa				
Prediction Function Order (2.1.3)	1st	2nd					
# of Prediction Functions per day (2.1.3.1)	1	3	6	12	24	48	96

2.2.1 Training/Ensemble Set Size

TESLA uses the ensemble database to train its prediction function(s). As the size of this ensemble set increases, the prediction quality also increases. Conversely, there also exists a minimum ensemble set size that depends on the other configuration parameters, below which the quality will drop significantly. For the TESLA configuration applied in this report, the minimum training set size was found to be 60 days, on average.

2.2.2 Number of Input Variables

TESLA produces a prediction function that can use any number of input variables. As the number of relevant input variables increases, the quality of the output also improves. However, the required minimum ensemble set size mentioned in the previous section also increases, and does so exponentially. It is important to find a balance between the number of input variables and the size of the available ensemble training data set.

The following WRF variables and groups of these variables were selected for training and prediction purposes as they are expected to have the greatest influence on observed GHI.

- All Variables: GHI, Columnar Cloud Cover (total cloud water mixing ratio in a vertical column), Surface Pressure, Geopotential Height of the 850 hPa and 700 hPa pressure levels,
  (Temperature, Relative Humidity, and u & v components of Wind Speed) at the surface, 850hPa, and 700hPa pressure levels, Surface Sensible Heat Flux, Latent Heat, Precipitation, Planetary Boundary Layer Height, Soil Temperature.
  In addition the Clear Sky Index was calculated based on WRF GHI and the Ineichen Clear Sky GHI Model.

- GHI, Columnar Cloud Cover, Surface Pressure, Height of 850 and 700hPa pressure levels

- GHI and Height of 700hPa pressure level

- GHI and Surface Pressure

- GHI and Columnar Cloud Cover

- GHI Only

- Surface Temperature, Temperature of 850 and 700hPa pressure levels

2.2.3 Prediction Function Order

The TESLA "order" indicates the order of the prediction function polynomial. Increasing the order improves the quality of the prediction function, but it also has multiple disadvantages. When the order is increased from 1 to 2 (i.e., from x, y, etc. to x, x^2, y, y^2, xy, etc.), the minimum number of required ensembles increases exponentially from about 60 to over 10,000. If a small training set were provided for training a high order TESLA, the forecast accuracy would degrade substantially. Due to restrictions in our training set (11,520 data points), we were only able to study orders of 1 and 2. If a large database existed, it would be possible to increase the order to improve the forecast.

Another disadvantage of increasing the order is the increase in computational complexity and cost for the training process. However, once the function is established, the prediction speed is essentially independent of the order.

2.2.4 Input of Observations and/or NWP output

While it is generally expected that the larger the ensemble set the more accurate the forecast, it is important to quantify the added value of including the NWP output versus just basing the analog method on past observations. For the input parameters we have considered three permutations:

- NWP output of GHI and observed GHI are input to TESLA.
- No NWP output is used. TESLA forecasts are only based on GHI observations.
- No temporally resolved input is used. TESLA forecasts reduce to a statistical average representation of a daily GHI cycle, i.e. only a bias correction is performed.

Only the GHI output of the NWP (versus the other variables listed in Section 2.2.2) was chosen for the first two permutations, because the input parameter comparison results in Section 4.4.

There are several options for including past observations as input parameters to the prediction function instead of or in addition to the WRF ensemble set. As an example, if the observation from 24 hours ago is used as an input parameter, TESLA will create a function that maps the observation from 24 hours ago into a prediction of current GHI. We can then use this function with our current observation to predict 24 hours into the future. This lag between the observation input and observation output is denoted by **D**.

It is possible to increase **D** beyond 24 hours to predict up to **D** hours in the future, but, the performance will decrease as **D** increases since current observation become less relevant to GHI far in the future. Note that, if an NWP output is used for the ensemble dataset, **D** must be less or equal to the forecast horizon of the ensemble dataset that provides the input parameters. If NWP is not used in the ensemble dataset, **D** is only limited by the number of observations available.

If **D** is decreased, the maximum forecast horizon will also decrease, but the output performance will increase significantly. Unless stated otherwise, **D** = 1 day to get a 24 hours ahead forecast.

2.2.5 Number of Prediction Functions

TESLA does not explicitly use the time of day information for its learning process. This allows us to create multiple prediction functions by dividing our ensemble database strategically. TESLA can use the whole training data set to produce a single prediction function that predicts all daylight hours of the day or it can divide the data set into (e.g. 24 different hour sets and produce one prediction function for each hour of the day 24 functions total).

Increasing the number of prediction functions improves the prediction quality as the functions can adapt to phenomena that tend to occur only during certain parts of the day. However, since the training set is divided into smaller sets, the effective size of the ensemble set for each function decreases. As mentioned before, this will decrease the forecast accuracy. It is important to find a balance between the number of prediction functions and the ensemble data set size.

3 Observations and Forecast Data

3.1 Observations

For validation, we used multiple measurement sites and forecast input parameters. The measurement locations and their abbreviations are shown in Table 2 and Figure 1. The sites contain complete weather instrumentation including measurements of GHI by a Licor Li200 photodiode pyranometer. The observational dataset has a temporal resolution of 10 minutes with an interval-ending time stamp and is available for 5 months (05/01/2013 – 09/30/2013). Observations with a positive GHI value during night time (12 a.m. to 4 a.m.) is assumed to be measurement error, thus removed.

Table 2: Weather stations with Global Horizontal Irradiance observations. Stations that are not used in the present report are marked in italics.

Abbreviation	Latitude	Longitude	Elevation (ft.)
CBD	33.14	-117.33	75
ESC	33.16	-117.03	725
FTV	33.26	-116.98	1833
RSF	33.03	-117.19	255
SPV	33.09	-116.96	430
HVW	33.07	-116.99	1142
LLC	33.26	-117.07	997
SOB	33.01	-117.28	15
MSD	32.81	-117.24	339
MTL	*32.84*	*-117.06*	*916*
RIO	32.84	-116.88	663
PSW1	33.12	-117.28	125
PSW3	32.83	-117.13	432
PSW4	33.14	-117.24	454
PSW5	32.83	-117.18	322
PSW6	33.13	-117.20	521
PSW7	32.84	-116.97	360
PSW8	32.86	-117.01	453

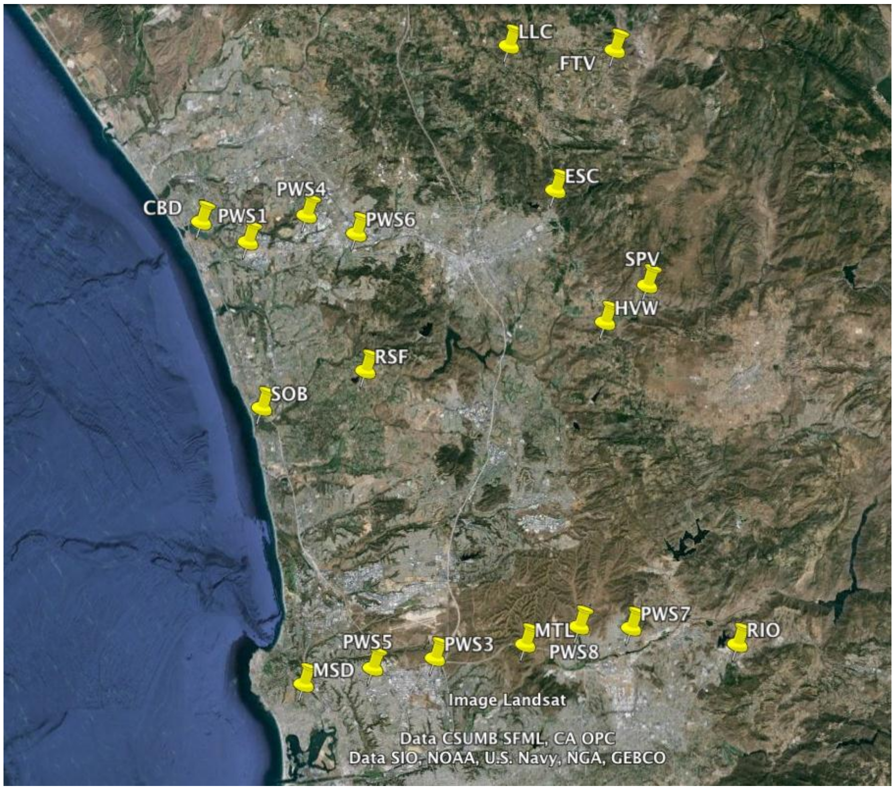


Figure 1: Markers denote the positions of the selected sites listed in Table 1. © Google Earth

3.2 Numerical weather prediction

As input parameters we have utilized GHI forecasts of different NWP products:

- GFS: Global Forecast System. Clear sky index interpolation was applied to generate a dataset with 15 min time steps.
- NAM: North American Mesoscale. Clear sky index interpolation was applied to generate a dataset with 15 min time steps.
- Green Power Labs (GPL) model. This is the probabilistic model developed under Task 3.1 of the same CSI contract and is described in more detail in a separate report by GPL[1]. The version of the probabilistic model applied here was an

[1] http://www.calsolarresearch.org/images/stories/documents/Sol3_funded_proj_docs/UCSD/ML-ProjRpt-CPUC_UCSD_2014-05-30.pdf

operational version during mid-phase of the project and is different from the final version applied in the GPL report (see "New Probabilistic" in the GPL report). The GPL report also uses a different geographic area and time period for validation and therefore the results are not directly comparable. Nevertheless, Fig. 17 in the GPL reports is consistent with Fig. 4 in this report in that the probabilistic model is biased towards forecasting clear conditions and has less than 10% hit rate in cloudy conditions for most forecast horizons.

- National Weather Service (NWS) post-processed NAM from the National Digital Forecast Database (NDFD)

- 24 hour persistence forecast

- UCSD post-processed Weather Research and Forecasting (WRF) model. Five different WRF model configurations were run. (see Section 3.2.1)

3.2.1 WRF: Weather Research and Forecasting

The WRF model is a state-of-the-art NWP and atmospheric simulation system that is maintained and supported as a community model (Skamarock et al., 2008). In this work, the version WRF V3.5 was used and configured with two nests of horizontal resolutions of 12.5 km and 2.5 km (Figure 2). To capture the characteristics of the shallow marine boundary layer, 75 terrain-following levels were used with 50 levels below 3000 m. The NAM, initialized at 12 UTC was used to derive boundary conditions for the outer domain. The NAM is run by the National Centers for Environmental Prediction (NCEP) and is based on WRF-NMM using 12 km horizontal resolution, 1-hour temporal resolution, and 60 vertical hybrid sigma-level coordinate. The WRF simulations were initialized at 0 UTC and run for 36 hours with 12 hours as spin-up time. The base WRF configuration, denoted as Base WRF (WRF without cloud data assimilation), is summarized in Table 3.

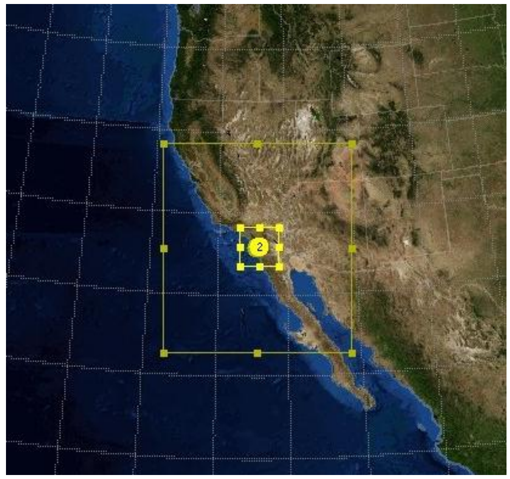


Figure 2: WRF simulation domains showing a nesting from a large domain with a spacing of Δx = 12.5 km to a small domain with a spacing of 2.5 km.

Table 3: Summary of WRF Configuration. For details please refer to the WRF user guide.

Domain/Time Options		**Physics Options (WRF Option #)**	
Δx (km)	2.5	Cumulus	NSAS (14)
Vertical Pts.	75	Radiation	New Goddard (5)
Output Interval (min)	15	Microphysics	Morrison (10)
Spin-Up (hr)	12	PBL	MYNN (5)
Initial & boundary conditions	12 UTC NAM	LSM	RUC (3)

Due to the difficulty of simulating marine layer stratocumulus, one configuration of WRF physics options is not able to consistently produce accurate forecast. Therefore, ensemble forecasts are created by running multiple forecasts, each with a unique variation in the configuration to represent the different sources of uncertainty. López-Coto et al., (2014) demonstrated that the cumulus scheme is the most important parameterization generating variability in simulating marine stratocumulus in coastal southern California. The second important parameterization is the radiation scheme. In addition, when the NSAS cumulus scheme was used, the microphysics option had a strong influence.

In addition to the model physics, the initial conditions were also varied. Mathiesen et al., (2013) developed the WRF-Cloud Data Assimilation (CLDDA) using Geostationary Operational Environmental Satellite (GOES) imagery to directly assimilate clouds in the initial conditions. Validated using the UCSD pyranometer network, the WRF-CLDDA was shown to be 17.4% less biased than the NAM.

Therefore, in addition to the base case listed in Table 3 , four WRF simulations were conducted using three cumulus, two radiation and two microphysics schemes. Table 4 showing the unique variations of each scheme relative to the base case.

The WRF ensemble datasets all have a temporal resolution of 15 minutes and are available for the same 5 months time frame as the observational dataset. Both datasets have a small number of missing data points due to problems in measurements and forecast simulation. In order to achieve a consistent data set of observations and WRF ensembles, all missing data points for a location have been removed from both WRF ensembles and observations. To temporally align observations and WRF forecasts, the station observations at :10 and :20 are averaged to match the :15 forecast data point and a similar procedure is used for the :45 data point. The :00 and :30 observations are already consistent with the WRF output times. These matching daytime datasets are then used for training and validation.

Table 4: Summary of unique configurations for four different WRF ensembles

Ensemble Name	Cumulus	Radiation	Microphysics
Cumulus1	Kain-Fritsch (1)	Dudhia (1) / RRTM (1)	
Microphysics8			Thompson (8) (8)
CLDDA			
CLDDA& Cumulus 3	Grell-Freitas (GF) (3)		

3.3 Raw Forecast Performance

The results of the raw (no postprocessing) forecast model output for the May – September marine layer forecast trials were compiled in Figure 3, **Error! Reference source not found.**, and **Error! Reference source not found.**. The NOAA models (NAM and GFS) have severe deficiencies in forecasting marine layer cloud cover; for the coastal CBD site in Fig. 3 all forecasts are clear. The National Weather Service (NDFD) post-processing correctly predicts morning clouds, but misses days that are completely overcast. The performance of the GPL model is the worst of any of the custom models with a hit rate on par with NAM and GFS. Persistence forecast and UCSD WRF configurations that include cloud data assimilation from satellite images (wrfcldda) perform best. In the San Diego summer climate, the absence of frontal passage causes significant 'inertia' in the weather conditions. Weather conditions change typically over time periods of 2-5 days and therefore a 24-hour persistence forecast is very accurate and difficult to beat.

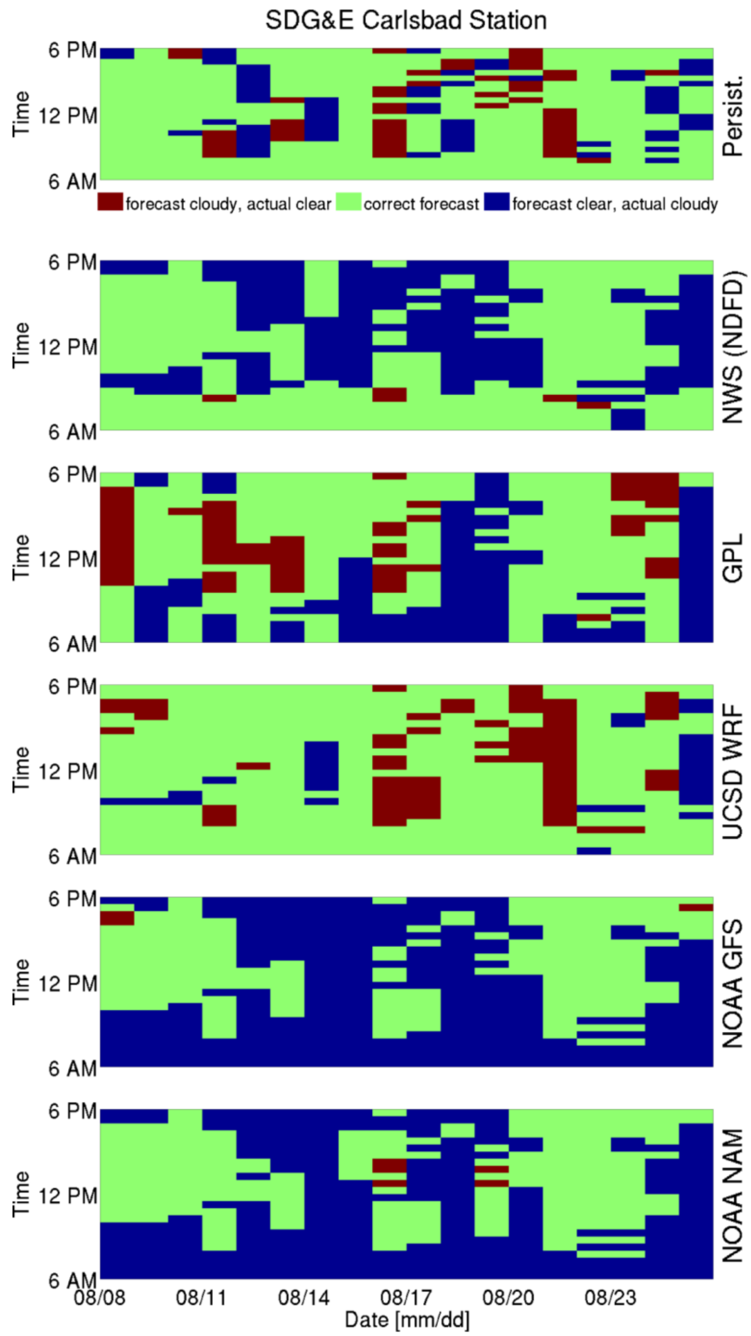


Figure 3: Forecast intercomparison between different solar forecast models relative to Global Horizontal Irradiation (GHI) measurements at Carlsbad, CA from August 8-25, 2013. Each column represents one day. Green color indicates a correct forecast. Blue color indicates a clear forecast that was actually cloudy, and red color indicates a cloudy forecast that was actually clear.

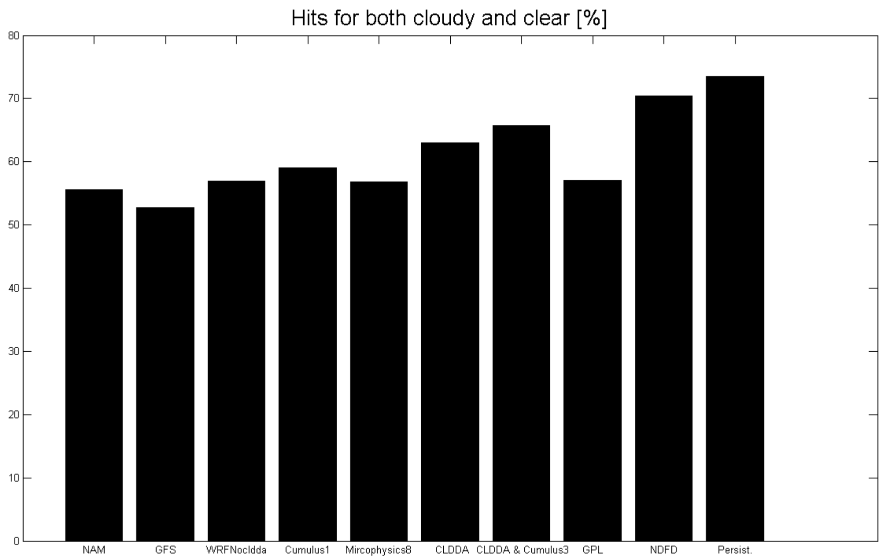


Figure 4: Hit score = 0.5 (Cloudy hit [%] + Clear hit [%]) for the forecast models averaged over CBD,.ESC, PWS1, PWS4, PWS5, PWS7 and PWS8 sites. The results at other stations are qualitatively similar.

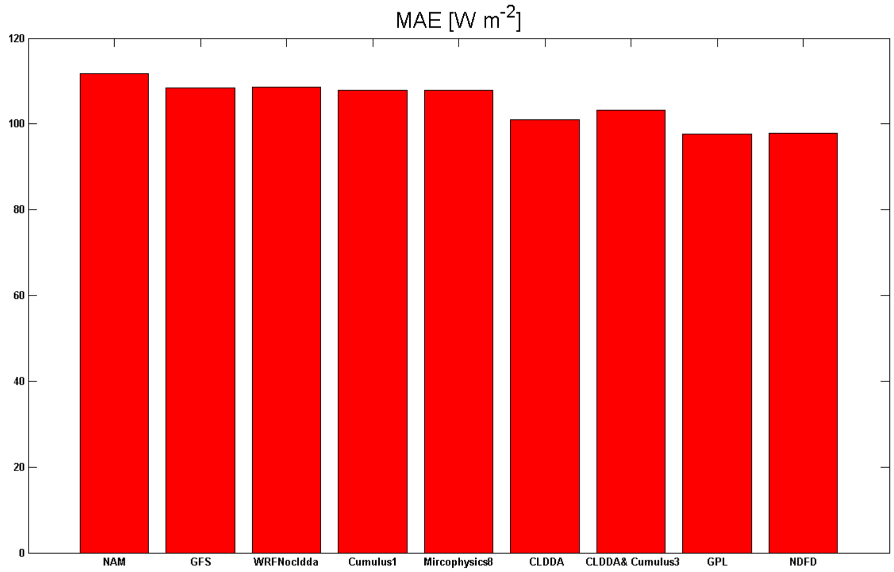


Figure 5: Daytime mean absolute error for the different weather forecast models.

3.4 Division of Dataset into Training and Verification

The dataset was divided into the training part and the verification/testing part. The training set size is maximized by using the entire dataset with the exception of the timeframe of 5 days after the forecast day. While using future data for training is unrealistic, this procedure results in a consistent number of training days for each forecast day. Removing the 5 immediately following days ensures that days that may be directly related to the forecast day are not used in the training.

The training dataset is used to obtain the forecast function for the selected day. Then, the WRF forecast variables for the selected day and station observations of the previous day are used to obtain the TESLA prediction, which is validated against the observation of the selected day. This procedure is repeated for all days within the dataset to obtain the overall performance that will be shown in the figures in Section 4.

3.5 Error metrics

Errors are computed for all 24 hours of the day. The main error metrics used within this report are Mean Absolute Error (MAE), Mean Bias Error (MBE) and Root Mean Square Error (RMSE) to quantify and compare the performance. Furthermore, another metric called forecast skill was used to quantify the relative performance between two forecasts models. MAE is obtained by averaging the ***absolute value*** of the error, to give a measure of the accuracy of the predictions

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Prediction_i - Observation_i|$$

MBE is obtained by taking the ***straight average*** of the error values to give an idea if the prediction tends to be systematically higher or lower than the observation,

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(Prediction_i - Observation_i)$$

RMSE is obtained by taking the average of the square of error values and taking the square root,

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Prediction_i - Observation_i)^2}$$

Forecast Skill measures the performance of TESLA with respect to the performance of the raw NWP (see Table 7 later) or with respect to the 24 hour persistence model (see Fig. 14 later). If there is no improvement, the forecast skill is 0. In the extreme case of no TESLA error, the forecast skill becomes 1.

$$\text{Forecast Skill} = 1 - \frac{RMSE_{Prediction}}{RMSE_{Reference}}$$

4 Post-processed forecast performance

4.1 Forecast Bias

A good post processing method is expected to remove bias and recover the average diurnal signal of GHI. Therefore as a first validation, the predictions are averaged over each hour of all days, providing an average diurnal cycle forecast for the CBD and ESC sites (Fig. 6). The coastal site (CBD) shows a typical marine layer signal with smaller irradiances in the morning. Figure 6 confirms that TESLA is essentially unbiased while the raw NWP forecast products are mostly positively biased and fail to produce morning clouds. Although all TESLA configurations perform similarly, only TESLA Order 2 results using NWP GHI is shown in the figures.

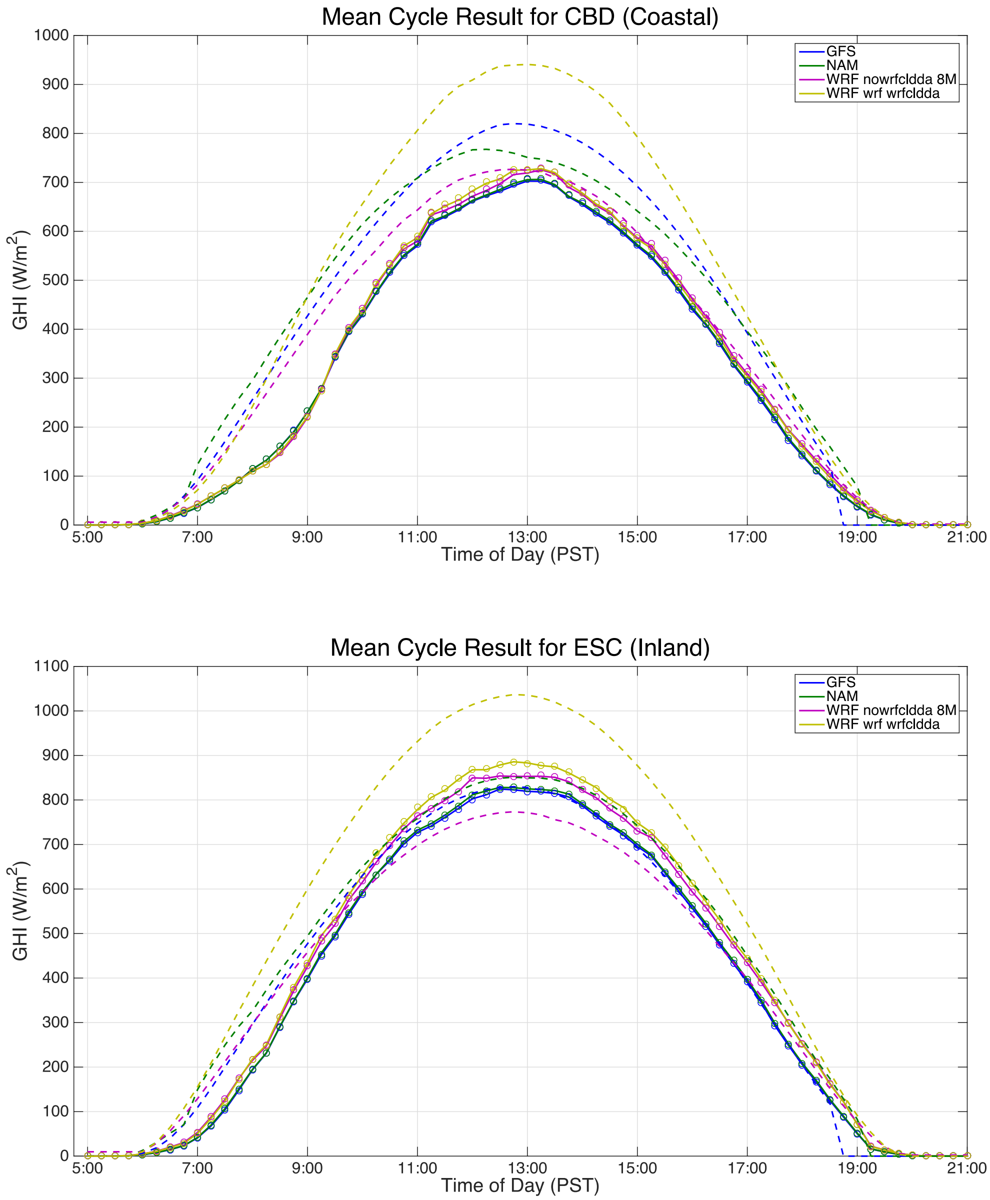


Figure 6: Average diurnal cycle for different forecasts over the May – September forecast period. The dashed lines show the 'raw' GHI forecast products of different Numerical Weather Prediction tools without post-processing. The solid lines show the TESLA Order 2 predictions using the GHI output of the forecast products. The circles show the observations. A) Coastal station CBD. (B) Inland station ESC.

The MBE and MAE results of the diurnal cycle figures are summarized in Table 5 and Table 6.

Table 5: Error Statistics for the CBD Coastal Station in W m^{-2}.

	CLDDA& Cumulus 3	CLDDA	Microphysics8	Cumulus	Base WRF	NAM	GFS
MAE Tesla	1.53	0.79	1.59	0.82	1.11	0.90	0.95
MAE Raw	54.00	85.70	24.40	89.23	82.43	58.25	61.32
MBE Tesla	-1.52	-0.60	-1.59	-0.58	-0.97	-0.85	-0.89
MBE Raw	53.90	85.58	23.30	88.60	82.30	57.44	58.57

Table 6: Error Statistics for the ESC Inland Station in W m-2.

	CLDDA &Cumulus3	CLDDA	Microphysics8	Cumulus	Base WRF	NAM	GFS
MAE Tesla	0.92	0.78	0.59	1.36	0.54	0.48	0.56
MAE Raw	66.11	65.77	31.72	47.05	45.46	31.7	16.84

4.2 Forecast RMSE: Choice of TESLA Configuration

TESLA forecast results for 120 days were analyzed to calculate error metrics and compare those to the RMSE of 24-hour persistence forecast. This section considers a separate function per every 15-minute interval resulting in 96 different functions. Nighttime prediction functions could be eliminated to reduce the computational cost, but are included within the error metrics.

First an optimal configuration of TESLA is analyzed in order to find the best order of the TESLA function and number of daily functions to include. In the following the TESLA input variables are varied. The forecast results are demonstrated only for the Carlsbad (CBD) station, but similar trends are observed at other stations. Two different TESLA methods are used:

- TESLA Order 1 applied to the ensemble set with past observations and NWP GHI
- TESLA Order 2 applied to the ensemble set with past observations and NWP GHI

Figure 7shows that the Order of the TESLA method has little impact on forecast error.

Furthermore, the number of prediction functions is varied as

- Creating a prediction function for each 15-minute interval of the day, resulting in 96 functions total.
- 30-minute intervals, resulting in 48 functions.
- 1-hour intervals, 24 functions.
- 2-hour intervals, 12 functions.
- 4-hour intervals, 6 functions.
- 8-hour intervals, 3 functions.
- 24-hour intervals, a single function.

Figure 7shows that the error decreases the finer the dataset is sliced into shorter time periods to generate the prediction functions. However, there is virtually no difference between 15 minute to 2 hour intervals. As intervals become shorter, the weather conditions at each time of day only vary slightly. A performance increase from more specificity may be offset by a reduction in the size of the training set.

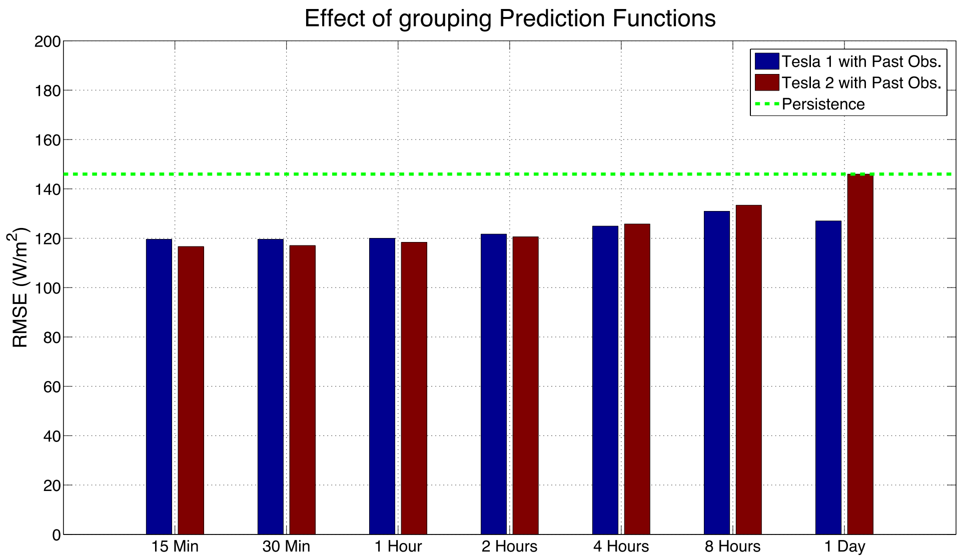


Figure 7: Effect of TESLA order and number of prediction functions on forecast accuracy. Creating more specialized prediction functions (one function every 15 minutes) decreases the error of the prediction.

For the remainder of this section only the following TESLA configuration of Order 2 with 15 minute prediction functions is used.

4.3 Sensitivity to type of input variables

In this section the forecast performance is examined for the scenarios described in Section 2.2.4. TESLA is applied (i) using no input (i.e. only a bias correction is performed), (ii) using NWP GHI only, (iii) past observations only, and (iv) using observations and NWP GHI as inputs. Table 7 shows the overall forecast skill and Figure 8,Figure 9,Figure 10 and Figure 11 show detailed results by station and WRF ensemble.

The forecast results are not strongly sensitive to the type of input variables. As expected, a bias correction (no input) performs worst and providing NWP and observation data results in the best forecast performance. However, the relatively persistent weather conditions in coastal California in the summer cause simple bias corrections to be quite effective. This performance difference may not be enough compared to the effort and time required to generate the WRF forecasts.

Table 7: Overall forecast results as a function of different input parameters and TESLA configurations. The forecast skill is measured with respect to the raw NWP forecast for various cases. A larger forecast skill is better.

	Overall Forecast Skill [1-RMSE(TESLA)/RMSE(NWP)]			
Input Type	None	NWP GHI Only	Past Observations Only	NWP GHI and Past Observations
TESLA Order **2** Func./15 Minute	0.44	0.47	0.48	**0.48**
TESLA Order **1** Func./15 Minute	0.44	0.45	0.47	0.48

Input Parameters (See Figure 12)	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**
TESLA Order **2** Func./15 Minute	-12.36	-0.86	0.36	0.37	0.03	**0.38**	0.34
TESLA Order **1** Func./15 Minute	0.00	0.36	0.36	0.37	0.35	0.36	0.36
TESLA Order **2** Func./1 Day	-2.03	0.18	0.22	0.21	0.21	0.23	0.20
TESLA Order **1** Func./1 Day	0.32	0.34	0.33	0.34	0.33	0.33	0.32

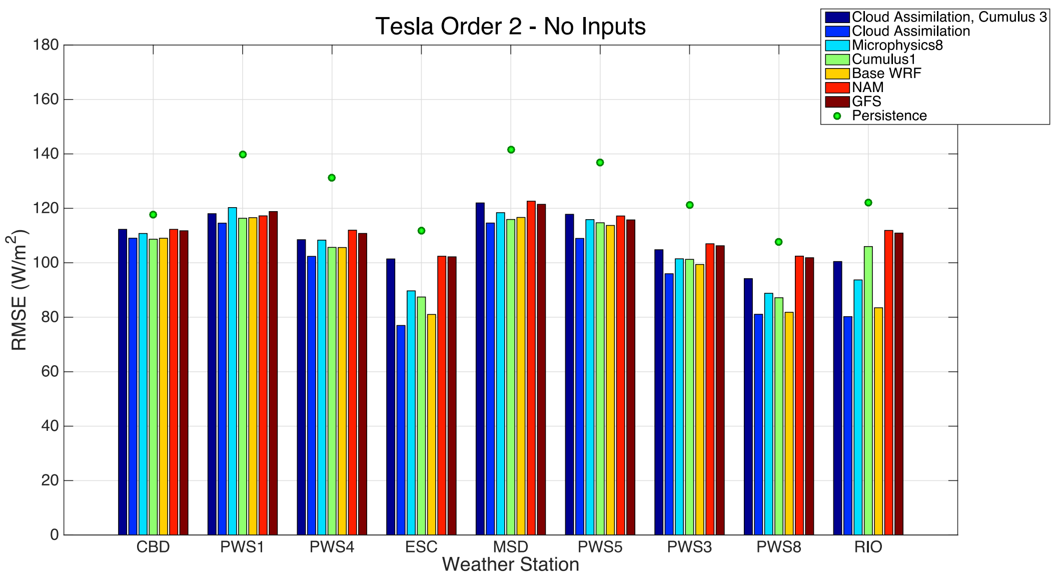


Figure 8: Marine layer solar forecast RMSE results with TESLA Order 2 without any input parameters. The prediction functions simply become constant values that correct the overall bias in the forecast.

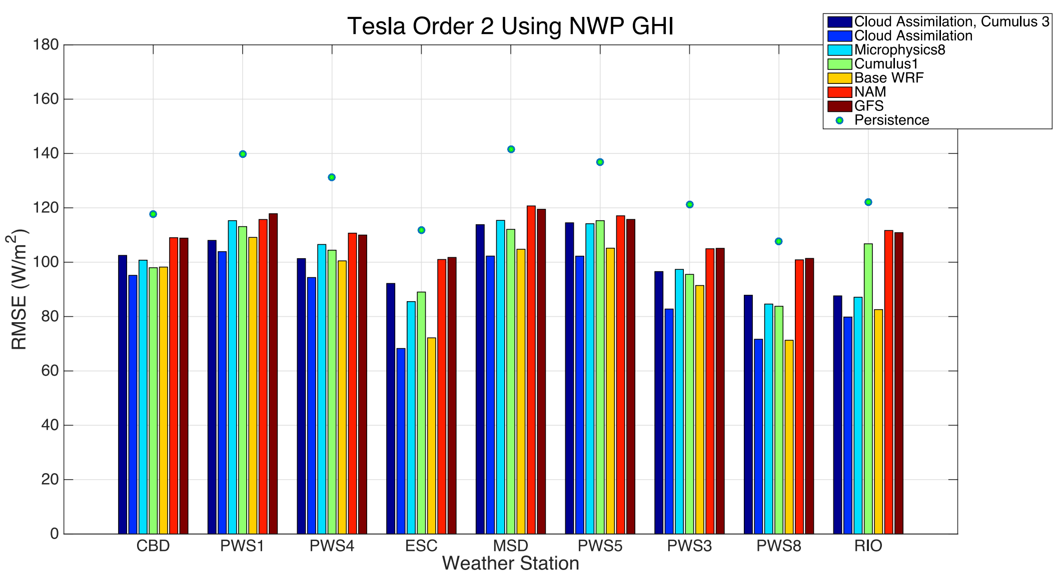


Figure 9: Marine layer solar forecast RMSE results with TESLA Order 2 using different NWP GHI outputs as input parameters, as shown in the legend.

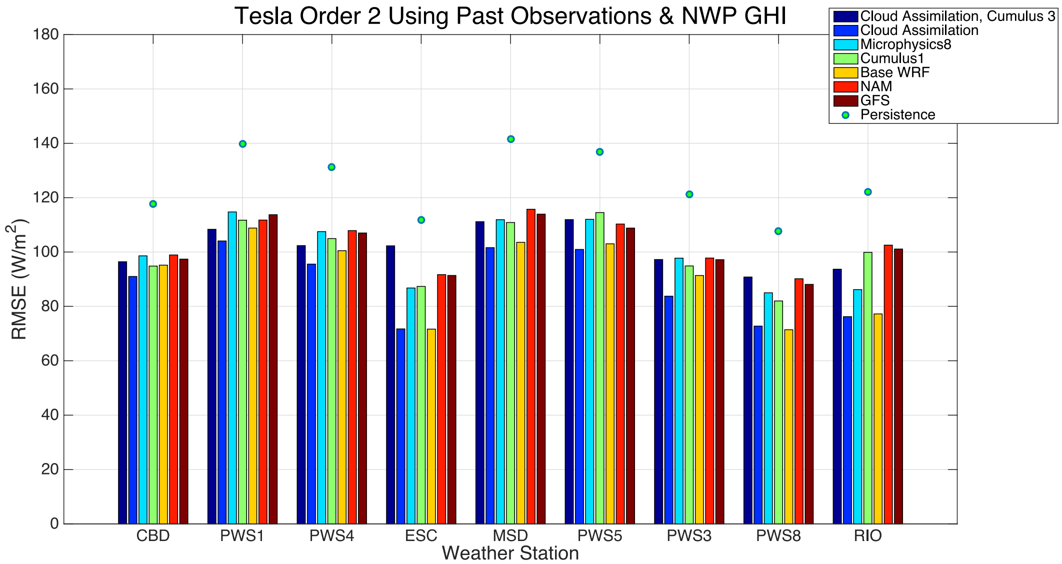


Figure 10: Marine layer solar forecast RMSE results with TESLA Order 2 using different NWP GHI outputs and past observations as input parameters.

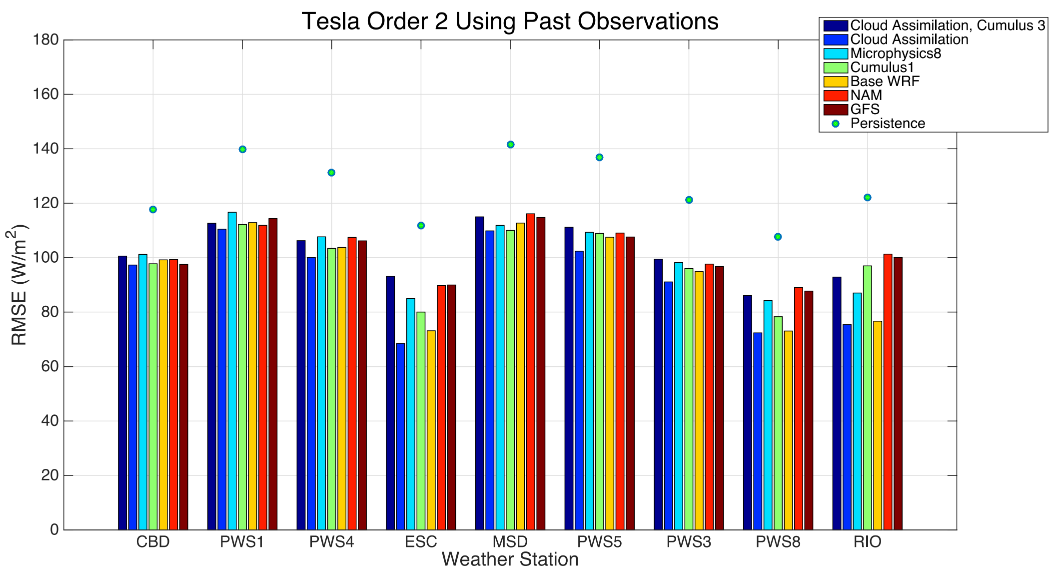


Figure 11: Marine layer solar forecast RMSE results with TESLA Order 2 using only past observations as input parameters.

4.4 Sensitivity to NWP input variables

In this section, we compared the effect of using different input parameters from NWP on the prediction quality. The WRF generated parameters given in Section 2.2.2 are used as prediction function inputs. Timeseries of all WRF generated inputs for a single example day are shown in the Appendix. For this analysis the base WRF configuration is used and TESLA output is compared against measurements at CBD.

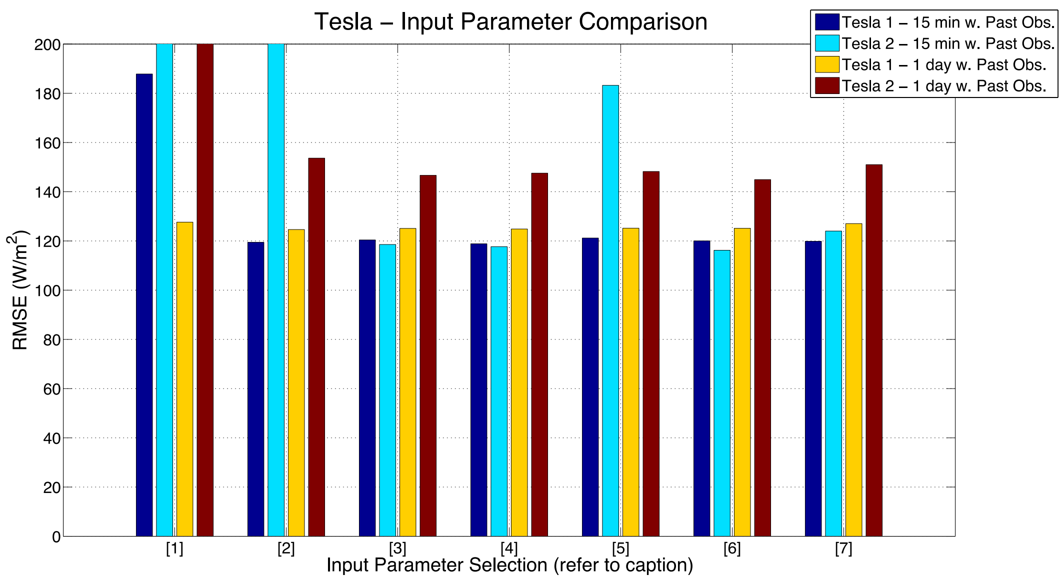
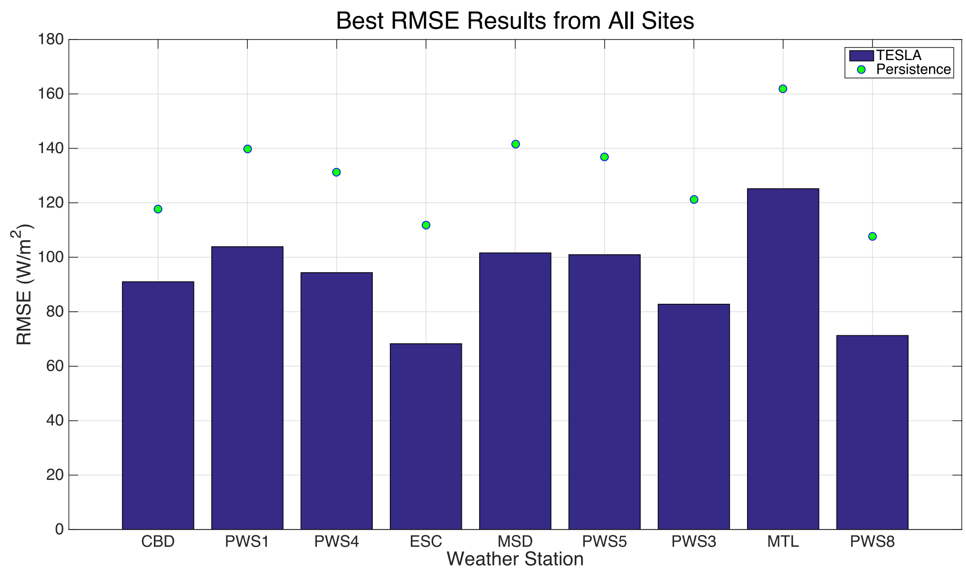


Figure 12: Marine layer solar forecast RMSE results with TESLA Orders 1 and 2 using different WRF output variables and past observations as input: [1] – All variables, [2] – GHI, Columnar Cloud Cover, Surface Pressure, Height 850 & 700, [3] – GHI and Height 700, [4] – GHI and Surface Pressure, [5] – GHI and Columnar Cloud Cover, [6] – GHI Only, [7] –Temperature at Surface, 850 & 700 mbar.

We can see that when all WRF variables are used to train the TESLA prediction function, large errors result, since the training dataset size requirement increases exponentially with increasing number of parameters. From the rest of the variable combinations, we conclude that a single WRF output - GHI – results in the smallest error for TESLA second order. Using a single variable is also the simplest to implement operationally.

4.5 Best TESLA Performance by Site

If the best result for every site is selected over all possible input parameters and forecast tools, the results in Fig. 12 are obtained. The improvement over 24 hour persistence ranges from 23% to 39%.



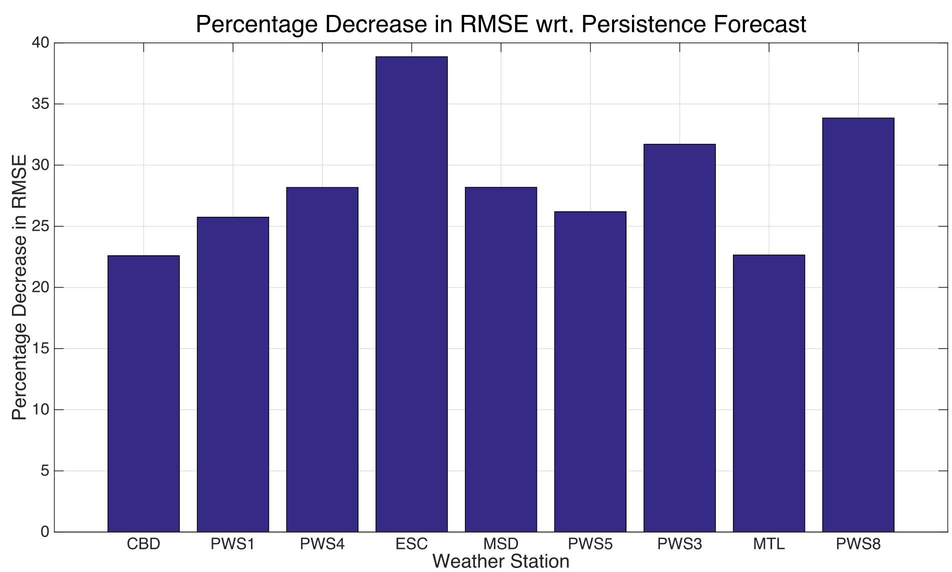


Figure 13: Best TESLA results for every site, selected over all possible input parameters and forecast tools shown in Figs. 7 through 10. a) RMSE. b) Forecast skill with respect to a 24 hour persistence forecast by the best performing TESLA implementation.

5 Conclusions

A solar forecast system with a focus on marine layer cloud forecasting in coastal Southern California during the summer time was developed. Several configuration of the Weather Research and Forecasting (WRF) model are applied to generate a distribution of forecast results for Global Horizontal Irradiance (GHI). The forecasts are validated against GHI observations form 8 weather stations located along a coastal to inland gradient.

The direct outputs from any numerical weather prediction model including WRF are significantly biased. Postprocessing is required to improve upon simple 24 hours persistence forecasts and for the WRF models this improvement is about 30% depending on the ground station. A postprocessing implementation that derives a different prediction function every 15 minutes and uses a second order algorithm delivers the best performance. The forecast performance is not sensitive to which input data (WRF output or past observations of GHI or both) are used. Also including additional WRF outputs aside from GHI shows little benefit.

The resulting forecast model is relatively easy to implement. Generating WRF forecasts is computationally expensive, but the postprocessing can be conducted quickly on a personal computer.

In the future, the forecasts will be implemented operationally on SDG&E computers. System-wide fleet and local feeder forecasts will present an opportunity for improved local voltage control and managing system-wide ancillary services to reduce operating costs and facilitate solar power integration.

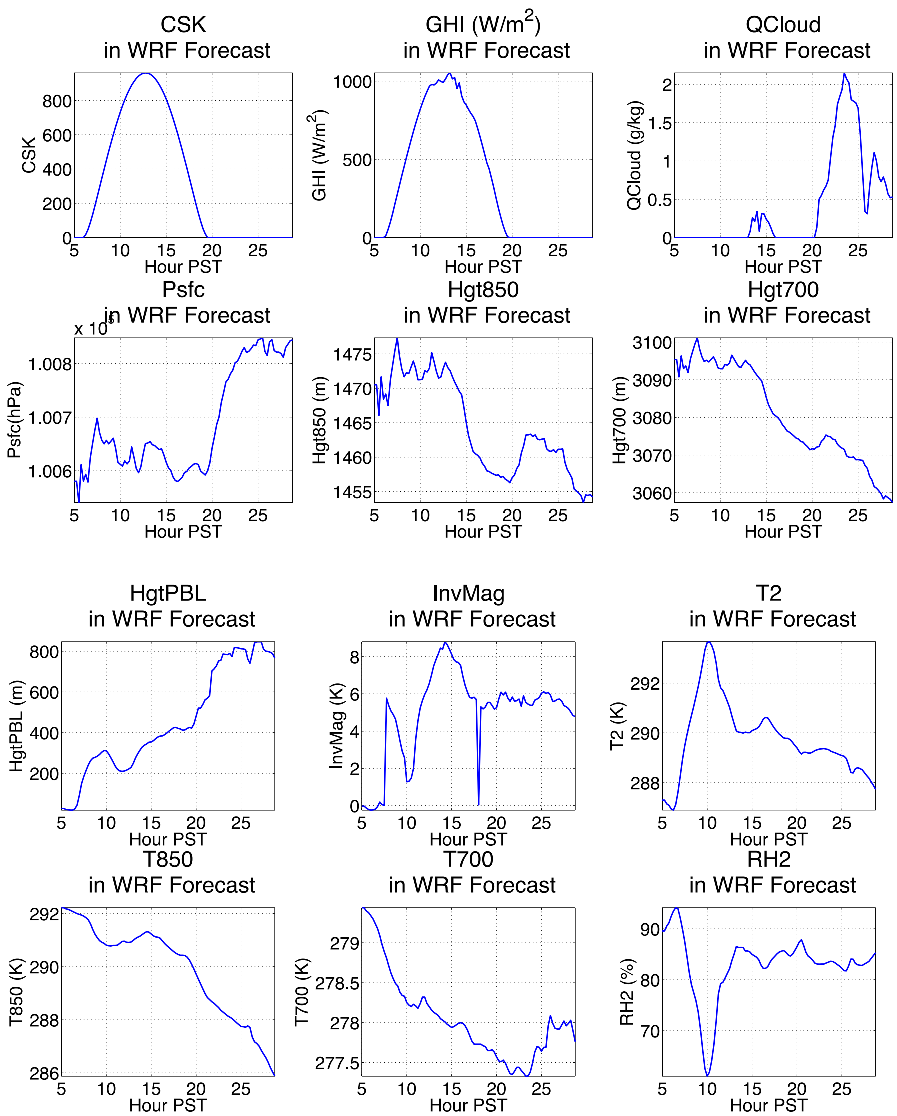
6 Works Cited

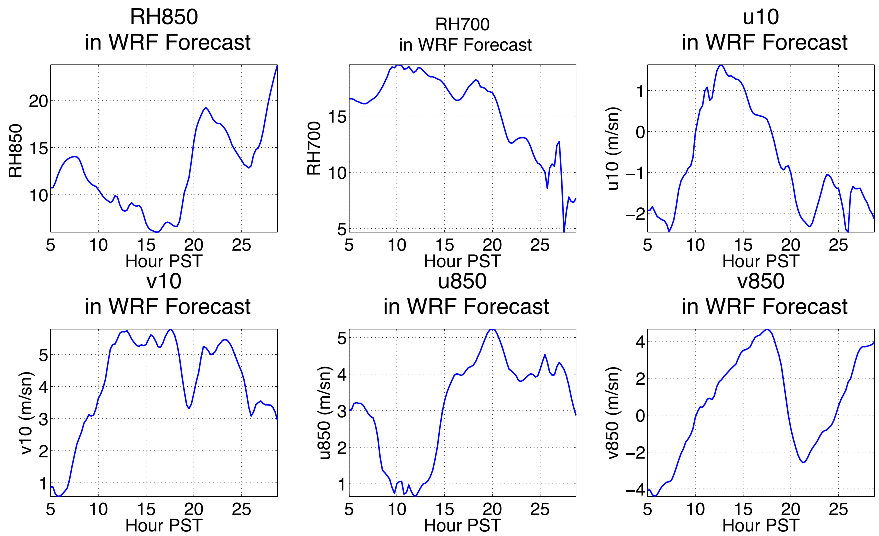
Maqsood, I. a. (2002). Intelligent weather monitoring systems using connectionist models. *NEURAL PARALLEL AND SCIENTIFIC COMPUTATIONS , 10*, 157-178.

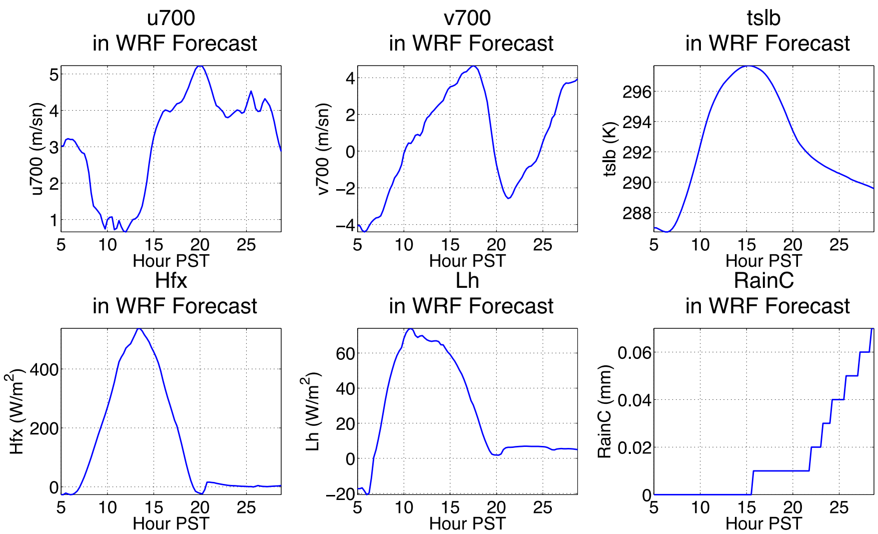
Mathiesen, PJ (2011). Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. *Solar Energy , 85*, 967-977.

Appendix A. Example NWP Input Timeseries

The following graphs show sample plots for the NWP inputs used in Section 4.4 for the example day of 05/01/12.




RH850
in WRF Forecast
RH850
Hour PST
RH700
in WRF Forecast
RH700
Hour PST
u10
in WRF Forecast
u10 (m/sn)
Hour PST
v10
in WRF Forecast
v10 (m/sn)
Hour PST
u850
in WRF Forecast
u850 (m/sn)
Hour PST
v850
in WRF Forecast
v850 (m/sn)
Hour PST


u700
in WRF Forecast
u700 (m/sn)
Hour PST
v700
in WRF Forecast
v700 (m/sn)
Hour PST
tslb
in WRF Forecast
tslb (K)
Hour PST
Hfx
in WRF Forecast
Hfx (W/m^2)
Hour PST
Lh
in WRF Forecast
Lh (W/m^2)
Hour PST
RainC
in WRF Forecast
RainC (mm)
Hour PST